(1) GENERAL INFORMATION:

   (i) APPLICANT: Reed, John

   (ii) TITLE OF INVENTION: Regulation of bcl-2 Gene Expression

  (iii) NUMBER OF SEQUENCES: 23

   (iv) CORRESPONDENCE ADDRESS:
      (A) ADDRESSEE: Gray, Cary, Ames & Frye
      (B) STREET: 401 B Street, Suite 1700
      (C) CITY: San Diego
      (D) STATE: CA
      (E) COUNTRY: USA
      (F) ZIP: 92101-4297

   (v) COMPUTER READABLE FORM:
      (A) MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
      (B) COMPUTER: IBM PC compatible
      (C) OPERATING SYSTEM: PC-DOS/MS-DOS
      (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

   (vi) CURRENT APPLICATION DATA:
      (A) APPLICATION NUMBER: US
      (B) FILING DATE:
      (C) CLASSIFICATION:

 (viii) ATTORNEY/AGENT INFORMATION:
      (A) NAME: Brotman, Harris F.
      (B) REGISTRATION NUMBER: 35461
      (C) REFERENCE/DOCKET NUMBER: P0041US0

   (ix) TELECOMMUNICATION INFORMATION:
      (A) TELEPHONE: (619) 699-3630
      (B) TELEFAX: (619) 236-1048

(2) INFORMATION FOR SEQ ID NO:1:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 20 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

   (iv) ANTI-SENSE: YES

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCCTTCCTAC CGCGTGCGAC

20

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 35 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CTTTTCCTCT GGGAAGGATG GCGCACGCTG GGAGA   35

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCTCCGACCC ATCCACGTAG   20

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 33 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACGGGGTACG GAGGCTGGGT AGGTGCATCT GGT   33

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTTGACGTCC TACGGAAACA   20

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 36 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCCCCAACTG CAGGATGCCT TTGTGGAACT GTACGG   36

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGGAAGGATG GCGCACGCTG   20

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGTGCGAC CCTCTTG   17

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TACCGCGTGC GACCCTC   17

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TCCTACCGCG TGCGACC   17

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid

```
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: DNA (genomic)

   (iv) ANTI-SENSE: YES


   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:                    17

CCTTCCTACC GCGTGCG

(2) INFORMATION FOR SEQ ID NO:12:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 17 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA (genomic)

     (iv) ANTI-SENSE: YES


     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:                  17

GACCCTTCCT ACCGCGT

(2) INFORMATION FOR SEQ ID NO:13:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 17 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA (genomic)

     (iv) ANTI-SENSE: YES

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:                  17

GGAGACCCTT CCTACCG

(2) INFORMATION FOR SEQ ID NO:14:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 15 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear
```

(ii) MOLECULE TYPE: DNA (genomic)

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGGCGGCAG CGCGG   15

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 15 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:   15

CGGCGGGGCG ACGGA

(2) INFORMATION FOR SEQ ID NO:16:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 16 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:   16

CGGGAGCGCG GCGGGC

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 18 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TCTCCCAGCG TGCGCCAT   18

(2) INFORMATION FOR SEQ ID NO:18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 18 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGCACTCACG CTCGGCCT   18

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5086 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
GCGCCCGCCC CTCCGCGCCG CCTGCCCGCC CGCCCGCCGC GCTCCCGCCC GCCGCTCTCC        60

GTGGCCCCGC CGCGCTGCCG CCGCCGCCGC TGCCAGCGAA GGTGCCGGGG CTCCGGGCCC       120

TCCCTGCCGG CGGCCGTCAG CGCTCGGAGC GAACTGCGCG ACGGGAGGTC CGGGAGGCGA       180

CCGTAGTCGC GCCGCCGCGC AGGACCAGGA GGAGGAGAAA GGGTGCGCAG CCCGGAGGCG       240

GGGTGCGCCG GTGGGGTGCA GCGGAAGAGG GGGTCCAGGG GGAGAACTT CGTAGCAGTC        300

ATCCTTTTTA GGAAAAGAGG GAAAAAATAA AACCCTCCCC CACCACCTCC TTCTCCCCAC       360

CCCTCGCCGC ACCACACACA GCGCGGGCTT CTAGCGCTCG GCACCGGCGG GCCAGGCGCG       420

TCCTGCCTTC ATTTATCCAG CAGCTTTTCG GAAAATGCAT TTGCTGTTCG GAGTTTAATC       480

AGAAGACGAT TCCTGCCTCC GTCCCCGGCT CCTTCATCGT CCCATCTCCC CTGTCTCTCT       540

CCTGGGGAGG CGTGAAGCGG TCCCGTGGAT AGAGATTCAT GCCTGTGTCC GCGCGTGTGT       600

GCGCGCGTAT AAATTGCCGA GAAGGGGAAA ACATCACAGG ACTTCTGCGA ATACCGGACT       660

GAAAATTGTA ATTCATCTGC CGCCGCCGCT GCCAAAAAAA AACTCGAGCT CTTGAGATCT       720

CCGGTTGGGA TTCCTGCGGA TTGACATTTC TGTGAAGCAG AAGTCTGGGA ATCGATCTGG       780

AAATCCTCCT AATTTTTACT CCCTCTCCCC CCGACTCCTG ATTCATTGGG AAGTTTCAAA       840

TCAGCTATAA CTGGAGAGTG CTGAAGATTG ATGGGATCGT TGCCTTATGC ATTTGTTTTG       900

GTTTTACAAA AAGGAAACTT GACAGAGGAT CATGCTGTAC TTAAAAAATA CAAGTAAGTC       960

TCGCACAGGA AATTGGTTTA ATGTAACTTT CAATGGAAAC CTTTGAGATT TTTTACTTAA      1020

AGTGCATTCG AGTAAATTTA ATTTCCAGGC AGCTTAATAC ATTGTTTTTA GCCGTGTTAC      1080

TTGTAGTGTG TATGCCCTGC TTTCACTCAG TGTGTACAGG GAAACGCACC TGATTTTTTA      1140

CTTATTAGTT TGTTTTTTCT TTAACCTTTC AGCATCACAG AGGAAGTAGA CTGATATTAA      1200

CAATACTTAC TAATAATAAC GTGCCTCATG AAATAAAGAT CCGAAAGGAA TTGGAATAAA      1260

AATTTCCTGC GTCTCATGCC AAGAGGGAAA CACCAGAATC AAGTGTTCCG CGTGATTGAA      1320
```

```
GACACCCCCT CGTCCAAGAA TGCAAAGCAC ATCCAATAAA ATAGCTGGAT TATAACTCCT    1380

CTTCTTTCTC TGGGGGCCGT GGGGTGGGAG CTGGGGCGAG AGGTGCCGTT GGCCCCCGTT    1440

GCTTTTCCTC TGGGAAGGAT GGCGCACGCT GGGAGAACGG GGTACGACAA CCGGGAGATA    1500

GTGATGAAGT ACATCCATTA TAAGCTGTCG CAGAGGGGCT ACGAGTGGGA TGCGGGAGAT    1560

GTGGGCGCCG CGCCCCCGGG GGCCGCCCCC GCACCGGGCA TCTTCTCCTC CCAGCCCGGG    1620

CACACGCCCC ATCCAGCCGC ATCCCGCGAC CCGGTCGCCA GGACCTCGCC GCTGCAGACC    1680

CCGGCTGCCC CCGGCGCCGC CGCGGGGCCT GCGCTCAGCC CGGTGCCACC TGTGGTCCAC    1740

CTGGCCCTCC GCCAAGCCGG CGACGACTTC TCCCGCCGCT ACCGCGGCGA CTTCGCCGAG    1800

ATGTCCAGCC AGCTGCACCT GACGCCCTTC ACCGCGCGGG GACGCTTTGC CACGGTGGTG    1860

GAGGAGCTCT TCAGGGACGG GGTGAACTGG GGGAGGATTG TGGCCTTCTT TGAGTTCGGT    1920

GGGGTCATGT GTGTGGAGAG CGTCAACCGG GAGATGTCGC CCCTGGTGGA CAACATCGCC    1980

CTGTGGATGA CTGAGTACCT GAACCGGCAC CTGCACACCT GGATCCAGGA TAACGGAGGC    2040

TGGGATGCCT TTGTGGAACT GTACGGCCCC AGCATGCGGC CTCTGTTTGA TTTCTCCTGG    2100

CTGTCTCTGA AGACTCTGCT CAGTTTGGCC CTGGTGGGAG CTTGCATCAC CCTGGGTGCC    2160

TATCTGAGCC ACAAGTGAAG TCAACATGCC TGCCCCAAAC AAATATGCAA AAGGTTCACT    2220

AAAGCAGTAG AAATAATATG CATTGTCAGT GATGTACCAT GAAACAAAGC TGCAGGCTGT    2280

TTAAGAAAAA ATAACACACA TATAAACATC ACACACACAG ACAGACACAC ACACACACAA    2340

CAATTAACAG TCTTCAGGCA AAACGTCGAA TCACCTATTT ACTGCCAAAG GGAAATATCA    2400

TTTATTTTTT ACATTATTAA GAAAAAAGAT TTATTTATTT AAGACAGTCC CATCAAAACT    2460

CCGTCTTTGG AAATCCGACC ACTAATTGCC AAACACCGCT TCGTGTGGCT CCACCTGGAT    2520

GTTCTGTGCC TGTAAACATA GATTCGCTTT CCATGTTGTT GGCCGGATCA CCATCTGAAG    2580

AGCAGACGGA TGGAAAAAGG ACCTGATCAT TGGGGAAGCT GGCTTTCTGG CTGCTGGAGG    2640

CTGGGGAGAA GGTGTTCATT CACTTGCATT TCTTTGCCCT GGGGGCGTGA TATTAACAGA    2700

GGGAGGGTTC CCGTGGGGGG AAGTCCATGC CTCCCTGGCC TCAAGAAGAG ACTCTTTGCA    2760

TATGACTCAC ATGATGCATA CCTGGTGGGA GGAAAGAGT TGGGAACTTC AGATGGACCT     2820

AGTACCCACT GAGATTTCCA CGCCGAAGGA CAGCGATGGG AAAAATGCCC TTAAATCATA    2880

GGAAAGTATT TTTTAAGCT ACCAATTGTG CCGAGAAAG CATTTTAGCA ATTTATACAA      2940

TATCATCCAG TACCTTAAAC CCTGATTGTG TATATTCATA TATTTTGGAT ACGCACCCCC    3000
```

```
CAACTCCCAA TACTGGCTCT GTCTGAGTAA GAAACAGAAT CCTCTGGAAC TTGAGGAAGT      3060

GAACATTTCG GTGACTTCCG ATCAGGAAGG CTAGAGTTAC CCAGAGCATC AGGCCGCCAC      3120

AAGTGCCTGC TTTTAGGAGA CCGAAGTCCG CAGAACCTAC CTGTGTCCCA GCTTGGAGGC      3180

CTGGTCCTGG AACTGAGCCG GGCCCTCACT GGCCTCCTCC AGGGATGATC AACAGGGTAG      3240

TGTGGTCTCC GAATGTCTGG AAGCTGATGG ATGGAGCTCA GAATTCCACT GTCAAGAAAG      3300

AGCAGTAGAG GGGTGTGGCT GGGCCTGTCA CCCTGGGGCC CTCCAGGTAG GCCCGTTTTC      3360

ACGTGGAGCA TAGGAGCCAC GACCCTTCTT AAGACATGTA TCACTGTAGA GGGAAGGAAC      3420

AGAGGCCCTG GGCCTTCCTA TCAGAAGGAC ATGGTGAAGG CTGGGAACGT GAGGAGAGGC      3480

AATGGCCACG GCCCATTTTG GCTGTAGCAC ATGGCACGTT GGCTGTGTGG CCTTGGCCAC      3540

CTGTGAGTTT AAAGCAAGGC TTTAAATGAC TTTGGAGAGG GTCACAAATC CTAAAAGAAG      3600

CATTGAAGTG AGGTGTCATG GATTAATTGA CCCCTGTCTA TGGAATTACA TGTAAAACAT      3660

TATCTTGTCA CTGTAGTTTG GTTTTATTTG AAAACCTGAC AAAAAAAAAG TTCCAGGTGT      3720

GGAATATGGG GGTTATCTGT ACATCCTGGG GCATTAAAAA AAAATCAATG GTGGGGAACT      3780

ATAAAGAAGT AACAAAAGAA GTGACATCTT CAGCAAATAA ACTAGGAAAT TTTTTTTTCT      3840

TCCAGTTTAG AATCAGCCTT GAAACATTGA TGGAATAACT CTGTGGCATT ATTGCATTAT      3900

ATACCATTTA TCTGTATTAA CTTTGGAATG TACTCTGTTC AATGTTTAAT GCTGTGGTTG      3960

ATATTTCGAA AGCTGCTTTA AAAAATACA TGCATCTCAG CGTTTTTTTG TTTTTAATTG      4020

TATTTAGTTA TGGCCTATAC ACTATTGTG AGCAAGGTG ATCGTTTTCT GTTTGAGATT      4080

TTTATCTCTT GATTCTTCAA AAGCATTCTG AGAAGGTGAG ATAAGCCCTG AGTCTCAGCT      4140

ACCTAAGAAA AACCTGGATG TCACTGGCCA CTGAGGAGCT TTGTTTCAAC CAAGTCATGT      4200

GCATTTCCAC GTCAACAGAA TTGTTTATTG TGACAGTTAT ATCTGTTGTC CCTTTGACCT      4260

TGTTTCTTGA AGGTTTCCTC GTCCCTGGGC AATTCCGCAT TTAATTCATG GTATTCAGGA      4320

TTACATGCAT GTTTGGTTAA ACCCATGAGA TTCATTCAGT TAAAAATCCA GATGGCGAAT      4380

GACCAGCAGA TTCAAATCTA TGGTGGTTTG ACCTTTAGAG AGTTGCTTTA CGTGGCCTGT      4440

TTCAACACAG ACCCACCCAG AGCCCTCCTG CCCTCCTTCC GCGGGGGCTT TCTCATGGCT      4500

GTCCTTCAGG GTCTTCCTGA AATGCAGTGG TCGTTACGCT CCACCAAGAA AGCAGGAAAC      4560

CTGTGGTATG AAGCCAGACC TCCCCGGCGG GCCTCAGGGA ACAGAATGAT CAGACCTTTG      4620

AATGATTCTA ATTTTTAAGC AAAATATTAT TTTATGAAAG GTTTACATTG TCAAAGTGAT      4680
```

```
GAATATGGAA TATCCAATCC TGTGCTGCTA TCCTGCCAAA ATCATTTTAA TGGAGTCAGT      4740

TTGCAGTATG CTCCACCTGG TAAGATCCTC CAAGCTGCTT TAGAAGTAAC AATGAAGAAC      4800

GTGGACGTTT TTAATATAAA GCCTGTTTTG TCTTTTGTTG TTGTTCAAAC GGGATTCACA      4860

GAGTATTTGA AAAATGTATA TATATTAAGA GGTCACGGGG GCTAATTGCT AGCTGGCTGC      4920

CTTTTGCTGT GGGGTTTTGT TACCTGGTTT TAATAACAGT AAATGTGCCC AGCCTCTTGG      4980

CCCCAGAACT GTACAGTATT GTGGCTGCAC TTGCTCTAAG AGTAGTTGAT GTTGCATTTT      5040

CCTTATTGTT AAAAACATGT TAGAAGCAAT GAATGTATAT AAAAGC                     5086
```

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 717 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 1..717

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
ATG GCG CAC GCT GGG AGA ACG GGG TAC GAC AAC CGG GAG ATA GTG ATG      48
Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
1               5                   10                  15

AAG TAC ATC CAT TAT AAG CTG TCG CAG AGG GGC TAC GAG TGG GAT GCG      96
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
                20                  25                  30

GGA GAT GTG GGC GCC GCG CCC CCG GGG GCC GCC CCC GCA CCG GGC ATC      144
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
            35                  40                  45

TTC TCC TCC CAG CCC GGG CAC ACG CCC CAT CCA GCC GCA TCC CGC GAC      192
Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
        50                  55                  60

CCG GTC GCC AGG ACC TCG CCG CTG CAG ACC CCG GCT GCC CCC GGC GCC      240
Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
65                  70                  75                  80

GCC GCG GGG CCT GCG CTC AGC CCG GTG CCA CCT GTG GTC CAC CTG GCC      288
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
                85                  90                  95

CTC CGC CAA GCC GGC GAC GAC TTC TCC CGC CGC TAC CGC GGC GAC TTC      336
Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
            100                 105                 110

GCC GAG ATG TCC AGC CAG CTG CAC CTG ACG CCC TTC ACC GCG CGG GGA      384
Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
        115                 120                 125

CGC TTT GCC ACG GTG GTG GAG GAG CTC TTC AGG GAC GGG GTG AAC TGG      432
Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
130                 135                 140
```

```
GGG AGG ATT GTG GCC TTC TTT GAG TTC GGT GGG GTC ATG TGT GTG GAG        480
Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145             150                 155                 160

AGC GTC AAC CGG GAG ATG TCG CCC CTG GTG GAC AAC ATC GCC CTG TGG        528
Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
                165                 170                 175

ATG ACT GAG TAC CTG AAC CGG CAC CTG CAC ACC TGG ATC CAG GAT AAC        576
Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
                180                 185                 190

GGA GGC TGG GAT GCC TTT GTG GAA CTG TAC GGC CCC AGC ATG CGG CCT        624
Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly Pro Ser Met Arg Pro
                195                 200                 205

CTG TTT GAT TTC TCC TGG CTG TCT CTG AAG ACT CTG CTC AGT TTG GCC        672
Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr Leu Leu Ser Leu Ala
210             215                 220

CTG GTG GGA GCT TGC ATC ACC CTG GGT GCC TAT CTG AGC CAC AAG            717
Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr Leu Ser His Lys
225             230                 235
```

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 239 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
 1               5                  10                      15

Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
            20              25                  30

Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
        35                  40                  45

Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
        50                  55                  60

Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
    65              70                  75                      80

Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
                85                  90                      95

Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
            100                 105                 110

Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
        115                 120                 125

Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
    130             135                 140

Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145             150                 155                     160

Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
            165                 170                 175

Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
            180                 185                 190

Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly Pro Ser Met Arg Pro
            195                 200                 205

Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr Leu Leu Ser Leu Ala
    210                 215                 220
```

Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr Leu Ser His Lys
225   230   235

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 615 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 1..615

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
ATG GCG CAC GCT GGG AGA ACG GGG TAC GAC AAC CGG GAG ATA GTG ATG    48
Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
 1               5                  10                  15

AAG TAC ATC CAT TAT AAG CTG TCG CAG AGG GGC TAC GAG TGG GAT GCG    96
Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
                 20                  25                  30

GGA GAT GTG GGC GCC GCG CCC CCG GGG GCC GCC CCC GCA CCG GGC ATC   144
Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
             35                  40                  45

TTC TCC TCC CAG CCC GGG CAC ACG CCC CAT CCA GCC GCA TCC CGC GAC   192
Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
         50                  55                  60

CCG GTC GCC AGG ACC TCG CCG CTG CAG ACC CCG GCT GCC CCC GGC GCC   240
Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
 65                  70                  75                  80

GCC GCG GGG CCT GCG CTC AGC CCG GTG CCA CCT GTG GTC CAC CTG GCC   288
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
                 85                  90                  95
```

```
CTC CGC CAA GCC GGC GAC GAC TTC TCC CGC CGC TAC CGC GGC GAC TTC       336
Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
            100             105                     110

GCC GAG ATG TCC AGC CAG CTG CAC CTG ACG CCC TTC ACC GCG CGG GGA       384
Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
        115             120                 125

CGC TTT GCC ACG GTG GTG GAG GAG CTC TTC AGG GAC GGG GTG AAC TGG       432
Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
        130             135                 140

GGG AGG ATT GTG GCC TTC TTT GAG TTC GGT GGG GTC ATG TGT GTG GAG       480
Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145             150                 155                 160

AGC GTC AAC CGG GAG ATG TCG CCC CTG GTG GAC AAC ATC GCC CTG TGG       528
Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
            165             170                 175

ATG ACT GAG TAC CTG AAC CGG CAC CTG CAC ACC TGG ATC CAG GAT AAC       576
Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
            180             185                 190

GGA GGC TGG GTA GGT GCA TCT GGT GAT GTG AGT CTG GGC                   615
Gly Gly Trp Val Gly Ala Ser Gly Asp Val Ser Leu Gly
        195             200             205
```

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 205 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Met Ala His Ala Gly Arg Thr Gly Tyr Asp Asn Arg Glu Ile Val Met
1               5                   10                  15

Lys Tyr Ile His Tyr Lys Leu Ser Gln Arg Gly Tyr Glu Trp Asp Ala
            20              25              30

Gly Asp Val Gly Ala Ala Pro Pro Gly Ala Ala Pro Ala Pro Gly Ile
        35              40              45

Phe Ser Ser Gln Pro Gly His Thr Pro His Pro Ala Ala Ser Arg Asp
    50              55              60

Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala
65              70              75              80
```

Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Ala
                85   90   95

Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Gly Asp Phe
            100   105   110

Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly
        115   120   125

Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp
    130   135   140

Gly Arg Ile Val Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu
145   150   155   160

Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
            165   170   175

Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn
            180   185   190

Gly Gly Trp Val Gly Ala Ser Gly Asp Val Ser Leu Gly
            195   200   205